DOCUMENT RESUME

ED 449 217 TM 032 362

AUTHOR VandenBosch, Dawn
TITLE Why Do We Factor Variables When We Care about Types of People? Q and Other Two-Mode Factor Analytic Models.
PUB DATE 2001-02-00
NOTE 26p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, February 1-3, 2001).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Factor Analysis; *Personality Assessment; *Q Methodology

ABSTRACT

This paper explains the types of entities that can be analyzed using two-mode factor analysis and the techniques associated with each variation. It also provides an in-depth review and example of one of the most popular variations, the Q-technique. Even when researchers have multiple variables, multiple people, and multiple occasions of measurement, one of the modes is typically held constant while performing a factor analysis on the other two modes, which results in the data collapsing onto a face of the "data box." In factor analysis, the matrix of associations is computed from a raw data matrix. The rows are replicates and the columns are the mode that is to be factored. There are six types of two-mode factor analytic techniques, designated R, Q, O, P, S, and T. The R-technique is the one traditionally thought of when discussing factor analysis. The O-technique and the P-technique hold the individual mode constant, and all the data are from one individual (or an average of all individuals). The S- and T-techniques are rarely used, and they are only generalizable to one variable. The Q-technique seeks to identify different types of people by factoring people over variables, holding occasions constant. The Q-technique often uses a type of ranking of responses known as the Q sort. The Q sort is explained, and a heuristic example is given of the use of Q-technique. (Contains 5 tables and 18 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Why do We Factor Variables When We Care About Types of People?

Q and Other Two-Mode Factor Analytic Models

Dawn VandenBosch

Texas A&M University 77843-4225

ED 449 217

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

D. Vandenbosch

1

TM032362

Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, February 1-3, 2001.

BEST COPY AVAILABLE

Factor analysis is a popular method of reducing the multiplicity of tests, variables or other entities (Kerlinger, 1986). For example, factor analysis is frequently used to discover the underlying latent variables that make up a certain construct. Thus, several factors such as verbal ability, mathematical ability, deductive reasoning may make up the construct of intelligence. Technically speaking, factor analysis attempts to find factors that are exist in a matrix of association such as a correlation matrix (Thompson, 1980). Developed by Spearman in 1904, the use of factor analysis increased with the advent of computers and statistical software (Thompson & Dennings, 1993).

Factor analysis is a procedure for analyzing two-dimensions or modes (Gorsuch, 1983). Traditionally, two-mode factor analysis uses a data matrix with the first mode including individuals and the second mode being variables presented as columns of the data matrix. There are methods that analyze three-modes however that information will not be covered in this paper. The interested reader is able to find more information in Gorsuch (1983).

The purpose of this paper is to explain the types of entities that can be analyzed using two-mode factor analysis and the techniques associated with each variation. Also, an in depth review and example of one of the most popular variations called the Q-technique will be discussed.

The Data Box

In 1946, Cattell proposed a covariation chart as a general analysis of the possible relations of choice in experimental studies which is now called the Basic Data Relation Matrix (Sells, 1963). Cattell (1966) discusses the Basic Data Relation Matrix to help explain how the three dimensions--individuals, variables and occasions--relate to each other. The box was originally used to demonstrated ten dimensions but because ten dimensions are difficult to score and analyze using factor analysis techniques, the box is now used to represent three modes(Cattell, 1966). From the box one can see three faces or surfaces of the box. The three surfaces merge into vectors. These vectors represent the three main modes. Three main correlation matrices can be transposed to equal six matrices that represent the six techniques of two-mode factor analysis.

As stated earlier, it is rare to use all three modes at the same time. Thus, even when researchers have multiple variables, multiple people, and multiple occasions of measurement, one of the modes is typically held constant while performing a factor analysis on the other two modes which results in the data collapsing onto a face (Cattell, 1966). The face is then represented by a raw data matrix which is used to create an estimate of association (Gorsuch, 1983). There are many different matrices of associations that can be used, but the most popular is the correlation matrix or Pearson r matrix (Carr, 1992).

Factor Analysis

In factor analysis the matrix of associations (e.g., correlation matrix) is computed from a raw data matrix (Thompson, 2000). The rows in the matrix are replicates and the columns are the the mode (i.e., participants, variables or occasions) which is to be factored. It is also important to remember that what differentiates the six techniques is the organization of the raw data matrix. However, it is erroneous to think one can transpose the data matrix to analyze the data by using another technique; this is because we want the number of rows to be several times larger than the number of columns.

Thompson (2000) noted

> …we generally want the number of rowrelationship/association pattern replications to be several times larger than the number of the column entities that we are factoring…This is to allow the patterns of relationships among the factored entities to be replicated over quite a number of rows in the raw data matrix, so that we can be sure that the estimated relationships are stable, and therefore that the factors we extract from the matrix of associations will themselves also be stable. (p. 9)

Thus, if we had data on 20 variables from 200 people, we could organize the people to constitute the rows of the data matrix while the columns represented the variables. A transposition of the data would result in a 20 row by 200 column

data matrix. Both forms of the data matrix are on the same surface on Cattell's data box, because the two modes for both organizations of the data involve multiple people and multiple variables measured only once. But the second organization of the data (i.e., 20 by 200) is not tenable for factor analysis, because stable estimates of patterns of relationships among the 200 people can not be estimated over only 20 row replications of relationship patterns.

For example, when using the Q-technique, Thompson (1981) noted that the number of variables as rows should be at least two more than twice the number of participants as columns. In a Q-technique then, you could have 67 variables then using this guideline, [(67-1)/2] = 33 subjects.

Now, we will discuss the six types two-mode factor analytic techniques. For a quick guide to determine which technique to use see Table 1.

INSERT TABLE 1 ABOUT HERE

The Six Two-Mode Techniques

The R and Q technique

There are six techniques of factor analysis: the R-technique, Q-technique, O-technique, P-technique, S-technique and T-technique. The technique that is chosen depends on what the researcher is interested in analyzing. Gorsuch (1983) noted that different factor analysis techniques seek to identify different

factors. The P- and R- technique seeks to identify states or traits. The O- and T- techniques seek to identify different situations or environments. And finally, the Q- and S- techniques try to identify different types of people. The discussion of techniques is organized here by a given technique and the transpose of the technique rather than their common factored modes.

The R-technique is the technique is what is traditionally thought of when discussing factor analysis. As stated earlier, the rows (replicates) are represented by individuals and the columns (mode being factored) represent variables. An example of this is the intelligence example given earlier. The transpose of this matrix is the Q-technique where rows are represented by variables and columns represented by participants. We will explore this further later in the paper. For both of these techniques, the mode that is held constant is the occasion.

The O and P techniques

The O-technique and the P-technique hold the individual mode constant and all data are from one individual (or an average of all individuals from each time and variable). The O-technique uses occasions as the factored mode and variables as the replicates. It seeks to identify the similarities in occasions for an individual (Gorsuch, 1983). For example, Jones, Thomson and Miller (1980) sought to simplify instructional choices for science educators. They used models of teaching as conditions or

occasions and indicators of instructional parameters as variables. They questioned which instructional parameters provide the basis for assigning the models to different categories (Jones et al., 1980). Jones et al, (1980) pooled the subjects' ratings of the instructional models (i.e., occasions) on the variables by computing the median rating of each occasion on each variable. A counseling example could use the same type of pooled rating for participants such as therapists to rate therapy models on variables such as certain therapy techniques to determine certain "super theories". Then, you can determine which techniques make up the certain factors or "super theories".

The transpose of the O-technique is the P-technique. The P-technique factors variables over occasions. Gorsuch (1983) considered the P-technique to be the forerunner to single-subject designs. Nesselroade and Ford (1985) consider the P-technique important in the study of intraindividual changes over occasions. The P-technique is a method used to show changes in scores for the same person at different points in time (Nunnaly, 1978). The technique factors variables over occasions with people being the mode that is held constant. Developmental psychologists may find this technique particularly useful. This technique can obtain a more rich and descriptive picture of a person's experience (Voelkl & Mathieu, 1995).

Mintz and Luborsky (1970) found the technique particularly useful when applied to the study of psychotherapy processes.

They used the method by looking at brief segments of verbal interchange from a total of nine sessions. They scored each segment on a variety of variables, then they intercorrelated the variables of the segments and finally factor analyzed the correlations (Mintz & Luborsky, 1970). The method would be very useful over different types of populations. For instance, one could measure depression over time in one individual to further understand how it develops or is maintained over different occasions. That information is useful in order to generalize or make inferences over that person's general status or well-being over time.

The S-technique and T-technique's are rarely used. There were no S-technique journal articles found in an exhaustive search over Psycinfo or ERIC documents. The reason for this according to Gorsuch (1983) is because the S and T techniques are only generalizeable to one variable. He suggests that because factor analysis comes from an area of psychology that is interested in individual differences, researchers that are interested in only one variable are not aware of these techniques. The S-technique factors individuals over occasions holding variables constant. The T-technique factors occasions over individuals holding variables constant. However, Frankiewicz and Thompson (1979) did report a T-technique factor analysis of types of teacher brinkmanship behaviors.

The Q-Technique

In 1917 Sir Cyril Burt proposed factoring people over a series of tests, though at the time he did not label the technique a "Q technique" (Cattell, 1978). In 1935, the British factorist, Sir Godfrey Thomson published a paper that outlined the possibilities of computing correlations between people rather than tests. Thomson named this technique "Q" to distinguish the technique from the traditional R technique. However, for various reasons, Thomson was pessimistic about Q-technique and did not pursue it further (Brown, 1980).

At almost the same time (independent of Thomson) William Stephenson was writing about the possibilities of person correlations and intrapersonal relationships (Brown, 1980; Stephenson, 1935). Stephenson introduced Q-methodology as a means of investigating human subjectivity. Stephenson elaborated on the theory and techniques of Q-methadology in his classic text, The study of behavior (Stephenson, 1953).

The Q-technique, as stated earlier, seeks to identify different types of people by factoring people over variables holding occasions constant. According to Thompson (2000) Q-technique seeks to answer three questions:

1. How many types of people are there?
2. Are the expected people associated with the expected factors?

3. Which variables were or were not useful in differentiating different types of people?

The Q-technique is an excellent technique for intensive study of a small number of people (Thompson, 2000). Technically, Q-technique is a way of rank-ordering variables and then assigning numbers to those ranks for statistical analysis (Kerlinger, 1986).

Q Sorts

Conventional-Sorting or Quasi-normal Method

The Q-technique often uses a type of ranking of responses called Q sorts. For example, five therapists may be asked to rank different items on a personality scale. The purpose of the Q-technique would be to see if there are different types of therapists in relation to their personality types. The therapist may be given 56 items of the Eysenck personality scale and asked to rank each item according to how much they are like this to how much they are not like this. The scale they would rank would be I am very much like this to I am not like this at all. The 56 items would each be statement is written on a different card. They would then take the cards and rank them. Kerlinger (1986) notes that usually, there are from 60 to 120 cards to sort. Furthermore, in a Q sort, he individual is asked to place the sorted cards in a quasi-normal distribution of ranks. Kerlinger (1986) uses the following example of a normal distribution of ranks with 90 cards:

Most Approve									Least Approve	
3	4	7	10	13	16	13	10	7	4	3
10	9	8	7	6	5	4	3	2	1	0

So, each person ranks the cards such that for Most Approve category 10 would have have 3 cards or items. For the middle, category 5, 16 cards would be in that pile. Thompson (1980) called this the "conventional-sorting method". This response format gives us data that are considered "ipsative" because it invokes a forced-choice response format in which the responses for one item force or limit the possible choices for the following items (Cattell, 1944; Thomspon, 2000). Ipsative data are useful in studying the common aspects in intraindividual differences such as in the Q-technique method (Thompson, 2000). This method is also useful because creates a quasi-normal distribution. Thompson (2000) noted:

> Thus, having data with exactly the same distributional shapes is appealing because when we correlate the participants, none of the person correlation coefficients will be attenuated by differences in score distribution shapes, even if we are computing a matrix of Pearson *r* coefficients as the basis for the Q-technique factor analysis. (p.10)

Thompson (1980) noted however, that this conventional-sorting method may be throwing away information. If a participant sorts five cards out in the Agree category they may not actually feel the same way towards every card in the category. Therefore,

you are losing the richness and value of that information. However, the task of rank ordering 90 items may be cognitively overwhelming (Thompson, 1980). Therefore, he suggested using a second type of sorting called the mediated-ranking method.

Mediated-Ranking Method

Another type of sorting is called mediated-sorting which makes the cognitively overwhelming task of rank ordering items easier (Thompson, 1980). Thompson (1980) proposed a two-stage method of sorting data with no ties. The partipants first rank order the cards in the conventional Q sort manner. Then, they go back to each category and rank order the cards within the piles. This method allows participants to rank order data with no ties without being so overwhelmed. Furthermore, this method allows for more variance and thus one is not losing as much information as with the conventional-sorting method (Thompson, 2000).

Unnumbered Graphic Scale

The prior two methods rank order data in a way that constrains participants to answering in a certain way. Furthermore, the prior two methods also lose some variance with regards to the data (Thompson, 1980). Thompson and Dennings (1993) noted that there is a third alternative to Q sorting normative data that allows for more variance, called the unnumbered graphic scale. A straight line is drawn between two antonyms such as "Agree" or "Disagree". The participants are asked to draw a mark on the line that best describes the position

of their answer to the item (Thompson & Dennings, 1993). The marks are scored using an equal-interval measurement scale with a large number of categories. Thompson (2000) noted that this method of Q sorting data mediates the cognitive complexity of ranking a large number of data with no ties while yielding more variable scores.

Thompson and Dennings (1980) also noted:

> Of course, using normative data will mean that the bivariate correlation coefficients analyzed in a Q-technique factor analysis will inherently be attenuated by variations in the distribution shapes of scores for different individuals, and that these differences will affect the identification of the factors extracted from the correlations…However, it is conceivable that tolerating some deviations in distribution shapes will not devastate the factor analytic solution, and may be worthwhile if not requiring people to make forced choices yields more accurate reflections of their feelings or beliefs. (p. 15)

Now that we have considered the types of Q sorting techniques used in a Q-technique factor analysis we can turn to the types of variables used.

Variables

Variables of any kind can be used with a Q-technique. They can be individual statements, pictures, or photographs (Thompson, 2000). There are actually two main ways of selecting these

variables: structured or unstructured responses. Structured responses are sampled from discrete item sets. For example, the Neo personality scale has over 100 items but those 100 items measure five different personality types. The variables themselves are implicitly structured (Thompson, 2000). Unstructued variables represent a single population of items (Thompson, 2000). For example if a researcher uses a scale that has no theoretical structure associated with it then it is unstructured. An example would be asking participants' opinions on empathetic behavior in therapy.

Heuristic Example

The following example will attempt to clarify the uses of a Q-technique. Let's say a researcher is interested in determining personality types among different types of psychologists. The researcher chooses those who work primarily in the academic arena (psychologist C, D, and E on Table 2) and those who work as clinicians (psychologist A and B in Table 2). The researcher then uses the mediated Q-sort technique where the psychologists rank order items on a personality scale. Table 2 shows the measure along with the ranking of the items by the psychologists. The psychologists ranked the personality items from "1" (most like me) to "25" (not like me at all).

INSERT TABLE 2 ABOUT HERE

Because the psychologists all ranked the items with no ties, the distribution statistics are identical. For our data set the mean equals 13 and the standard deviation is 7.36 for all five psychologists. This is a convenient and efficient way to check for data entry errors in your research (Thompson, 2000).

Table 3 presents the correlation matrix for the data set. It is fairly easy to see that there are two factors just by looking at the correlation matrix. It is also easy to see that the two factors separate the academic psychologists from the psychologists who identify themselves as clinicians. This is just a heuristic example; a correlation matrix with a larger sample would be less clear. Thompson (2000) noted that in real practice we estimate person factors empirically rather than just by looking at a bivariate correlation matrix.

INSERT TABLE 3 ABOUT HERE

The "pattern/structure coefficients" are presented in Table 4. The table presents the two person factors extracted from the correlation matrix. "Each pattern/ structure coefficient represents the correlation of a given person with a given person factor" (Thompson, 2000, p. 18). In this example, psychologist D is most correlated with person Factor I (r_s = .928). This table answers the question, "How many types of people are there?"

INSERT TABLE 4 ABOUT HERE

Finally, Table 5 presents the factor scores associated with each factor. Often absolute 1.0 is used as a cutscore to interpret Q-technique factor scores (Thomspon, 2000). In Table 5 the three variables most useful for defining person Factor I were variables 17 (1.72932), 10 (1.69836) and 3 (-1.51889). The three most useful for defining person Factor II were variables 18 (1.62095), 22 (-1.51134) and 1 (-1.45644). Table 5 then answers the question "Which variables were and were not useful in differentiating the various person types/factors"? Factor scores can be thought as a ranking of the variables as regards the persons defining a given person factor (Thompson, 2000).

INSERT TABLE 5 ABOUT HERE

Table 5 also represents the items (or personality features in this example) that are most important to each type of psychologist. Remember, 1 was described as the "most like me" and 25 was represented as "most unlike me". Therefore, items 10 (-1.69836) and 3 (-1.51889) were most like the academics and items 1 (-1.45644) and 22 (-1.51134) were most like the clinicians. Items 17 (1.72932) and 21 (1.43746) were most unlike the academics and items 18 (1.62095) and 13 (1.42913) were most unlike the clinicians.

Hopefully, by this example one can clearly see the benefits of using the Q-technique to distinguish different types of people. For a more in depth view and understanding of the Q-

technique, see Thompson (2000), Gorsuch (1983), and Kerlinger (1986).

Factor analysis is very useful when one wants to reduce the multiplicity of tests and variables. Hopefully, I have shown that it is also useful when using situations and people also. Also, it is important to note that one must use the correct technique for the research you are conducting. When we wish to study types of people, factoring variables is not appropriate.

References

Brown, S. R. (1980). Political subjectivity. New Haven: Yale University Press.

Carr, S. C. (1992). A primer on the use of Q-technique factor analysis. Measurement and Evaluation in Counseling and Development, 25, 133-138.

Cattell, R. B. (1944). Psychological measurement: Normative, ipsative, interactive. Psychological Review, 51, 292-303.

Cattell, R. B. (1966). The data box: Its ordering of total resources in terms of possible relational systems. In R. B. Cattell (Ed.), Handbook of Multivariate Experimental Psychology (pp. 67-128). Chicago: Rand McNally.

Frankiewicz, R. G., & Thompson, B. (1979, April). A comparison of strategies for measuring teacher brinkmanship behavior. Paper presented at the annual meeting of the American Educational Research Association, San Fransisco. (ERIC Document Reproduction Service No. ED 171 753)

Gorsuch, R. L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Erlbaum.

Jones, H. L., Thompson, B., & Miller, A. H. (1980). How teachers perceive similarities and differences among various teaching models. Journal of Research in Science Teaching, 17, 321-326.

Kerlinger, F. N. (1986). Foundations of behavioral research (3rd ed.). New York: Holt, Reinhart and Winston.

Mintz, J. & Luborsky, L. (1970). P-technique factor analysis in psychotherapy research: An illustration of a method. Psychotherapy, Theory, Research, and Practice, 7, 13-18.

Nesselroade, J. R. & Ford, D. H. (1985). P-technique comes of age. Research on Aging, 7, 46-80.

Nunnaly, J. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Spearman, C. (1904). "General intelligence," objectively determined and measured. American Journal of Psychology, 15, 201-293.

Stephenson, W. (1935). Technique of factor analysis. Nature, 136, 297-310.

Stephenson, W. (1953). The study of behavior. Chicago: The University of Chicago Press.

Thompson, B. (1980). Validity of an evaluator typology. Educational Evaluation and Policy Analysis, 2, 59-65.

Thompson, B. (2000). Q-technique factor analysis: One variation on the two-mode factor analysis of variables. In L. Grimm & P. Yarnold (Eds.), Reading and understanding more multivariate statistics (pp. 207-226). Washington, DC: American Psychological Association.

Thompson, B. & Dennings, B. (1993, November). The unnumbered graphic scale as a data-collection method: An investigation

comparing three measurement strategies in the context of Q-technique factor analysis. Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 364 589)

Voelkl, J. E. & Mathieu, M. A. (1995). Intraindividual variation in the subjective experiences of older adults residing in a nursing home. Therapeutic Recreation Journal, 2, 115-123.

Table 1
Six Two-mode Factor Analytic Techniques

Technique	Factored Mode	Replicated Mode	Constant Mode
O	Occasions	Variables	Individuals
P	Variables	Occasions	Individuals
Q	Participants	Variables	Occasions
R	Variables	Participants	Occasions
S	Participants	Occasions	Variables
T	Occasions	Participants	Variables

Table 2

Data for Q-technique Psychologist Personality Study

		Psychologists				
	Items	A	B	C	D	E
1	Is talkative	1	2	23	24	20
2	Tends to find fault with others	4	5	22	23	19
3	Does a thorough job	15	16	2	3	1
4	Is depressed...blue	22	19	5	10	9
5	Is original	12	13	3	1	4
6	Is reserved	19	21	6	7	7
7	Is helpful and unselfish with others	23	22	12	11	12
8	Can be somewhat careless	16	17	21	21	21
9	Is relaxed, handles stress well	11	11	16	17	18
10	Is curious about many different things	5	4	1	2	3
11	Is full of energy	10	14	19	18	17
12	Is a reliable worker	9	9	7	6	8
13	Can be tense	24	24	8	5	5
14	Is ingenious, a deep thinker	8	7	4	4	2
15	Generates a lot of enthusiasm	20	20	13	12	13
16	Has a forgiving nature	14	12	20	20	22
17	Tends to be disorganized	17	18	25	25	24
18	Tends to be quiet	25	25	9	9	10
19	Likes to reflect, play with ideas7	8	10	8	11	
20	Can be cold and aloof	21	23	11	13	14
21	Is sometimes shy, inhibited	13	15	24	22	23
22	Is considerate and kind to almost everyone	2	1	15	19	25
23	Prefers work that is routine	6	6	18	14	15
24	Is outgoing, sociable	3	3	17	15	16
25	Is sometimes rude to others	18	10	14	16	6

Table 3

Correlation Matrix for Q-technique Data

	Psychologist				
Person	PsyA	PsyB	PsyC	PsyD	PsyE
PsyA	1.000				
PsyB	.955	1.000			
PsyC	-.238	-.188	1.000		
PsyD	-.221	-.200	.962	1.000	
PsyE	-.262	-.188	.902	.918	1.000

Table 4

Varimax-rotated Pattern/Structure Coefficient Matrix

	Factor	
Person	I	II
PsyA	-.137	.980
PsyB	-8.6E	.985
PsyC	.973	-.106
PsyD	.979	-.103
PsyE	.956	-.121

Table 5

Sorted Factor Scores on Factor I and II from the Factor Score Matrix

Item	Factor Score I	Factor Score II
1	1.14677	-1.45644
2	1.05239	-1.05230
3	-1.51889	.17525
4	-.59515	.96778
5	-1.47520	-.23823
6	-.78556	.87936
7	-.04055	1.30809
8	1.19224	.61970
9	.53494	-.21705
10	-1.69836	-1.36974
11	.70014	-.05601
12	-.91639	-.65862
13	-.81841	1.42913
14	-1.45884	-.92473
15	.06307	.97454
16	1.08385	.12039
15	.06307	.97454
16	1.08385	.12039
17	1.72932	.81982
18	-.33103	1.62095
19	-.56094	-.82590
20	.09665	1.25486
21	1.43746	.30192
22	.75665	-1.51134
23	.26771	-.93732
24	.26696	-1.35285
25	-.12886	.12873

BEST COPY AVAILABLE